A Minimalist Parsing Analysis of Subject/Object Asymmetries in Basque Relative Clauses

Overview A top-down parser for Minimalist grammars (MGs; Stabler 2013) has been used in recent years to successfully model sentence processing preferences across an array of languages and phenomena, when combined with complexity metrics that relate parsing behavior to memory usage (Gerth 2015; Kobele et al. 2013:a.o.). This approach provides a quantifiable theory of the effects of detailed grammatical structure on cognitive cost, and can help strengthen the link between generative syntactic theory and sentence processing. In this sense, subject/object relative clause (RC) asymmetries have been extensively probed with this model for a variety of languages (Graf et al. 2017; De Santo 2021), partially due to their popularity in the psycholinguistic literature more broadly. Here, we contribute to this line of work by extending these results to the analysis of relative clauses in Basque, as the syntactic properties of this language set it apart from other languages studied so far in the existing MG processing literature.

MG Parsing and Memory Metrics. The MG model links structural details to processing difficulty by associating the stack states of a (deterministic) top-down parser (Stabler 2013) to memory burden. In the context of this abstract, memory usage (Gibson 2000; Kobele et al. 2013) is mainly measured based on how long a node is kept in memory (*tenure*). Consider the MG derivation in Fig. 1. The index of a node n encodes the moment n was predicted and put in memory by the parser. The outdex encodes the moment n is confirmed and flushed out of memory. Tenure for n is measured as $outdex(n) - index(n)$: e.g. *tenure*(*does*) $= 8 - 3 = 5$ (Fig. 1). For derivations including long-distance dependencies, an additional measure of cost (*size*) evaluates the length of movement and its interaction with the details of the encompassing structure. These notions are then implemented in a set of **off-line** metrics of processing difficulty (i.e., MaxT and SumSize) over a whole derivation, and used to compute categorical pairwise comparisons (Graf et al. 2017).

Processing Asymmetries for Basque RCs. Basque is an ergative and head-final language with both pre- and post-nominal RCs (de Rijk 2007), making it an ideal candidate to test the MG model. Sentence processing experiments show a general preference for subject (SRC, 1a) over object (ORC, 1b) RCs (Munarriz et al. 2016). As the MG parser is sensitive to fine-grained grammatical structure, details of the syntactic choices are crucial. We consider two syntactic analyses for Basque RCs, adapted from similar approaches often employed cross-linguistically: a head internal analysis (Gondra 2016:HI), and a head external one (Artiagoitia 1992:HE). In a nuthsell, the former proposes that an external determiner selects the RC whose head moves from its base subject/object position to Spec, CP (its landing site within the RC) and further operations produce the correct order. The HE analysis posits the presence of an RC-internal null operator coindexed with an external DP (the head of the RC). It is this null operator that raises to Spec, CP.

Modeling Results and Discussion. Following Munarriz et al. (2016), we tested postnominal RCs like in (1) for syntactic derivations corresponding to the two analyses considered. Modeling results show the parser correctly predicts a preference for SRCs over ORCs for both analyses, but with important differences. For the HE analysis, MaxT leads to a tie between SRCs and ORCs, then ORCs turn out to be more complex due to additional movement dependencies. In the HI case, MaxT makes the correct prediction directly, due to the need to wait for the object extraposition to head of the RC. These results are in line with previous work, and support the idea that the MG model offers ways of probing the impact of subtle structural choices. Finally, psycholinguistic work has pointed out that in specific morpho-syntactic environments, properties related to ergativity can lead Basque to show an ORC preference (Carreiras et al. 2010). The MG approach will allow us to further explore these claims with reference to particular syntactic theories, by isolating properties related to the case system in unmarked structures.

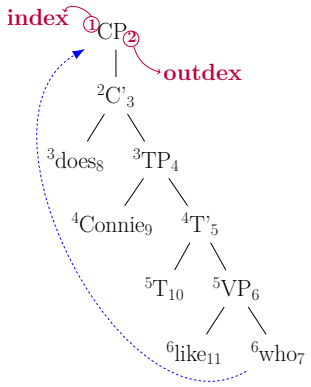

Figure 1: Example of an MG derivation tree with parse steps.

	Head Internal		Head External	
	MaxT	**Node**	**MaxT**	**Node**
SRC	30	orain	25	orain
ORC	31	orain	25	orain

Table 1: Performance of MaxT for each of the RC derivations, according to a Head Internal or Head External analysis.

(1) Example of test sentences

a. Irakasleak aipatu ditu-en ikasleak lagunak ditu orain.
teacher.pl.abs mention.prt has=comp student.sg.erg friend.pl.abs has now
'The student that mentioned the teachers has friends now.' **SRC**

b. Irakasleak aipatu ditu-en ikasleak lagunak dira orain.
teacher.pl.abs mention.prt has=comp student.pl.erg friend.pl.abs are now
'The students that the teachers mentioned are friends now.' **ORC**

Artiagoitia, X. 1992. Why basque doesn't relative everything? *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*. • Carreiras, M., Jon Andoni Dunabeitia, M. V., de la Cruz-Pavia, I., and Laka I. 2010. Subject relative clauses are not universally easier to process: Evidence from Basque. *Cognition*. • De Santo, A. 2021. Italian postverbal subjects from a minimalist parsing perspective. *Lingue e linguaggio*. • Gerth, S. 2015. *Memory limitations in sentence comprehension: A structural-based complexity metric of processing difficulty*. PhD Thesis. Universitatsverlag Potsdam. • Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain*. • Graf T., J. Monette, and C. Zhang. 2017. Relative clauses as a benchmark for Minimalist parsing. *J. of Language Modelling.* • Kobele, G.M., S. Gerth, and J. Hale. 2013 Memory resource allocation in top-down minimalist parsing. *Formal Grammar*. • Munarriz, A., Ezeizabarrena, M.J., Gutierrez-Mangado, M.J. 2016. Differential and selective morpho-syntactic impairment in spanish-basque bilingual aphasia. *Bilingualism: Language and Cognition*. • de Rijk, R. P.G. 2007. *Standard Basque: A progressive grammar*. The MIT Press. • Stabler, E. P. (2013). Two models of minimalist, incremental syntactic analysis. *Topics in cognitive science*.